HEALPix: A FRAMEWORK FOR HIGH-RESOLUTION DISCRETIZATION AND FAST ANALYSIS OF DATA DISTRIBUTED ON THE SPHERE

K. M. Górski, ^{1,2} E. Hivon, ³ A. J. Banday, ⁴ B. D. Wandelt, ^{5,6} F. K. Hansen, ⁷ M. Reinecke, ⁴ and M. Bartelmann ⁸ Received 2004 September 21; accepted 2004 December 10

ABSTRACT

HEALPix—the Hierarchical Equal Area isoLatitude Pixelization—is a versatile structure for the pixelization of data on the sphere. An associated library of computational algorithms and visualization software supports fast scientific applications executable directly on discretized spherical maps generated from very large volumes of astronomical data. Originally developed to address the data processing and analysis needs of the present generation of cosmic microwave background experiments (e.g., BOOMERANG, WMAP), HEALPix can be expanded to meet many of the profound challenges that will arise in confrontation with the observational output of future missions and experiments, including, e.g., Planck, Herschel, SAFIR, and the Beyond Einstein inflation probe. In this paper we consider the requirements and implementation constraints on a framework that simultaneously enables an efficient discretization with associated hierarchical indexation and fast analysis/synthesis of functions defined on the sphere. We demonstrate how these are explicitly satisfied by HEALPix.

Subject headings: cosmic microwave background — cosmology: observations — methods: statistical

1. INTRODUCTION

Advanced detectors in modern astronomy generate data at huge rates over many wavelengths. Of particular interest to us are those data sets that accumulate measurements distributed on the entire sky, or a considerable fraction thereof. Typical examples include radio, cosmic microwave background (CMB), submillimeter, infrared, X-ray, and gamma-ray sky maps of diffuse emission, and full-sky or wide-area surveys of extragalactic objects. Together with this wealth of gathered information comes an inevitable increase in complexity for data reduction and science extraction. In this paper we are focused on those issues related to the distinctive nature of the spherical spatial domain over which the data reside. Our original motivations arose from work related to the measurement and interpretation of the CMB anisotropy. The growing complexity of the associated science extraction problem can be illustrated by the transition between the data sets from various experiments: COBE Differential Microwave Radiometer (DMR) (early 1990s, 7° FWHM resolution, \sim 6000 pixel sky maps at three wavelengths), BOOMERANG (late 1990s, 12' FWHM, partial sky maps of $\sim 200,000$ pixels at four wavelengths), WMAP (early 2000s, resolution up to 14' FWHM, \sim 3 million pixel sky maps at five wavelengths), and Planck (data expected ca. 2009, resolution up to 5' FWHM, \sim 50 million pixel sky maps at nine wavelengths). Science extraction from these data sets involves the following:

- 1 JPL/Caltech, MS 169-327, 4800 Oak Grove Drive, Pasadena, CA 91109. 2 Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warsaw, Poland.
- ³ IPAC, MS 100-22, Caltech, 1200 East California Boulevard, Pasadena, CA 91125.
- ⁴ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Strasse 1, Postfach 1317, D-85741 Garching bei München, Germany.
 - ⁵ Department of Physics, University of Illinois, Urbana, IL 61801.
- ⁶ Department of Astronomy, University of Illinois, 1002 West Green Street, Urbana, IL 61801.
- ⁷ Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, N-0315 Oslo, Norway.
- 8 ITA, Universität Heidelberg, Tiergartenstrasse 15, D-69121 Heidelberg, Germany.

- 1. global analysis problems: harmonic decomposition, estimation of the power spectrum, and higher order measures of spatial correlations;
- 2. real space morphological analyses, object detection, identification, and characterization;
- 3. the simulation of models of the primary and foreground sky signals to study instrument performance and calibrate foreground separation and statistical inference methods; and
- 4. spatial and/or spectral cross-correlation with external data sets.

These tasks, and many others, necessitate a careful definition of the data models and proper construction of the mathematical framework for data analysis such that the algorithmic and computing time requirements can be satisfied in order to achieve the successful and timely scientific interpretation of the observations. A particular method of addressing some of these issues is described next.

2. DISCRETIZED MAPPING AND ANALYSIS OF FUNCTIONS ON THE SPHERE

The analysis of functions on domains with a spherical topology occupies a central place in both the physical sciences and engineering disciplines. This is particularly apparent in the fields of astronomy, cosmology, geophysics, and atomic and nuclear physics. In many cases the geometry is dictated either by the object under study or by the need to assume and exploit approximate spherical symmetry to utilize powerful perturbative techniques. Practical limits for the purely analytical study of these problems create an urgent necessity for efficient and accurate numerical tools.

The simplicity of the spherical form belies the intricacy of global analysis on the sphere. There is no known point set that achieves the analog of uniform sampling in Euclidean space and allows exact and invertible discrete spherical harmonic decompositions of arbitrary but band-limited functions. All existing practical schemes proposed for the treatment of such discretized functions on the sphere introduce some (hopefully

small) systematic errors dependent on the global properties of the point set. The goal is to minimize these errors and faithfully represent deterministic functions as well as realizations of random variates (both in configuration and Fourier space) while maintaining computational efficiency.

We illustrate these points using as an example the field of CMB anisotropies. Here we are already witnessing an explosive growth in the volume of the available data. Indeed, full-sky measurements of the CMB anisotropy in both total intensity and polarization by current and future generations of surveys present serious challenges to those involved in the analysis and scientific exploitation of the results.

A pixelized sky map is an essential intermediate, and often highly compressed, stage in the processing of data between the raw form acquired by instrumental detectors and the final stage of analysis resulting in estimates of typically a few values for the physical parameters of interest. The COBE DMR sky maps comprising 6144 pixels per map were considered large at the time of their release (ca. 1992). As for both the currently available (e.g., WMAP) and forthcoming (e.g., ESA's Planck satellite) CMB data sets, the associated full-sky surveys at an angular resolution of $\sim 10'$ (FWHM) should be comprised of a few pixels per resolution element (so that the discretization effects on the signal are subdominant with respect to the effects of the instrument's angular response) and will therefore require map sizes of at least $N_{\rm pix}$ of the order of a few times 1.5×10^6 pixels. Even more pixels will be needed to fully render the angular resolution of the Planck HFI high-frequency channels. This estimate, $N_{\rm pix}$, should be multiplied by the number of frequency bands (or indeed, by the number of individual observing channels—74 in the case of *Planck*—to account for the analysis work performed before optimal co-added maps are made for each frequency band) to yield the approximate expected size of this compressed form of the survey data, which is then the input to the astrophysical analysis pipeline. Clearly, careful attention must be given to devising data structures that can maximally facilitate the efficient analyses of such large data sets, especially given that many essential scientific questions can only be answered by studies of their global properties.

This paper describes the essential geometric and algebraic properties of our method for the digital representation of functions on the sphere—the Hierarchical Equal Area and isoLatitude Pixelization (HEALPix)—and the associated multipurpose computer software package. This software has been made available to the community since 1997.

3. REQUIREMENTS FOR A SPHERICAL PIXELIZATION SCHEME

The numerical analysis of functions on the sphere involves (1) a class of mathematical operations, whose objects are (2) discretized maps, i.e., quantizations of arbitrary functions according to a chosen tessellation (an exhaustive partition of the sphere into finite area elements). Hereafter, although we mostly specialize our discussion to the CMB-related applications of HEALPix, our statements generally hold true for any relevant set of deterministic and random functions on the sphere.

Considering point 1: The standard operations of numerical analysis that one might wish to execute on the sphere include convolutions with local and global kernels, Fourier analysis with spherical harmonics and power spectrum estimation, wavelet

decomposition, nearest neighbor searches, topological analyses including searches for extrema or zero crossings and the computation of Minkowski functionals, and the extraction of patches and finite differencing for solving partial differential equations. Some of these operations become prohibitively slow if the sampling of functions on the sphere and the related structure of the discrete data set are not designed carefully.

Regarding point 2: Typically, a full-sky map rendered by a CMB experiment contains

- (a) signals coming from the sky, which are by design strongly bandwidth-limited (in the sense of spatial Fourier decomposition) by the instrument's angular response function;
- (b) a projection into the elements of a discrete map, or pixels, of the observing instrument's noise; this pixel noise should ideally be random, and white, at least near the discretization scale, with a bandwidth significantly exceeding that of all the signals.

With these considerations in mind we proposed the following list of desiderata for the mathematical structure of discrete full-sky maps:

- 1. Hierarchical structure of the database. This is recognized as essential for very large databases and was indeed postulated already in the construction of the Quadrilateralized Spherical Cube¹⁰ (hereafter QuadCube; see White & Stemwedel 1992), which was used for the *COBE* sky maps. A simple argument in favor of this states that the data elements that are nearby in a multidimensional configuration space (here, on the surface of a sphere) are also nearby in the tree structure of the database. This property facilitates various topological methods of analysis and allows for easy construction of wavelet transforms on triangular and quadrilateral grids through fast lookup of nearest neighbors.
- 2. Equal areas for the discrete elements of the partition. This is advantageous because white noise at the sampling frequency of the instrument gets integrated exactly into white noise in the pixel space, and sky signals are sampled without regional dependence (although care must be taken to choose a pixel size sufficiently small compared to the instrumental resolution to avoid pixel shape—dependent signal smoothing).
- 3. Isolatitude distribution for the discrete area elements on the sphere. This property is essential for computational speed in all operations involving evaluations of spherical harmonics. Since the associated Legendre polynomials are evaluated via slow recursions, any sampling grid deviations from an isolatitude distribution result in a prohibitive loss of computational performance with the growing number of sampling points.

Various alternatives for sampling distributions on the sphere have been employed for the discretization and analysis of functions (e.g., see Driscoll & Healy 1994, Muciaccia et al. 1998 [rectangular grids]; Baumgardner & Frederickson 1985, Tegmark 1996 [icosahedral grids]; Saff & Kuijlaars 1997, Crittenden & Turok 1998 ["igloo" grids]; Szalay & Brunner 1999 [a triangular grid]), but each fails to meet simultaneously all the above requirements. In particular:

- (a) The QuadCube obeys desiderata 1 and (approximately) 2 but fails on desideratum 3 and cannot be used for efficient Fourier analysis at high resolution.
- (b) The Equidistant Cylindrical Projection (ECP), a very common computational tool in geophysics and climate modeling, recently implemented for CMB work by Muciaccia et al. (1998),

⁹ HEALPix is currently distributed from the Web site http://www.eso.org/science/healpix. However, in the near future the HEALPix Web site will move to the new address http://healpix.jpl.nasa.gov.

See http://lambda.gsfc.nasa.gov/product/cobe/skymap_info_new.cfm.

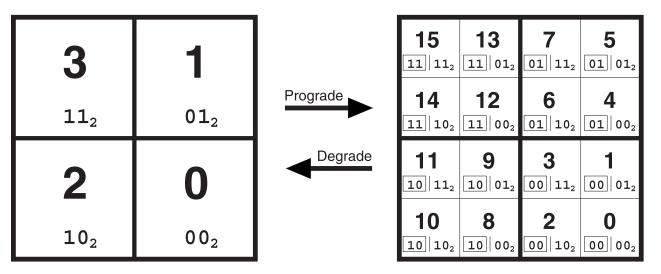


Fig. 1.—Quadrilateral tree pixel numbering scheme. The coarsely pixelized coordinate patch on the left consists of 4 pixels. Two bits suffice to label the pixels. To increase the resolution, every pixel splits into 4 daughter pixels, shown on the right. These daughters inherit the pixel index of their parent (boxed) and acquire two new bits to form the new pixel index. Several such curvilinearly mapped coordinate patches (12 in the case of HEALPix, and 6 in the case of the COBE QuadCube) are joined at the boundaries to cover the sphere. All pixel indices carry a prefix (here omitted for clarity) that identifies which base-resolution pixel they belong to.

satisfies desiderata 1 and 3 but by construction fails with desideratum 2. This is a nuisance from the point of view of application to full-sky survey data as a result of wasteful oversampling near the poles of the map. While the angular resolution of the measurements is fixed by the instrument and does not vary over the sky, the map resolution, or pixel size, depends on the distance from the poles. This must also be accounted for in work related to the integration of data or discretized functions over the sphere.

- (c) Hexagonal sampling grids with icosahedral symmetry perform superbly in those applications where near uniformity of sampling on the sphere is essential (Saff & Kuijlaars 1997), and they can be devised to meet desideratum 2 (Tegmark 1996). However, by construction they fail both desiderata 1 and 3.
- (d) Igloo-type constructions are devised to satisfy desideratum 3 (E. L. Wright 1997, private communication; Crittenden & Turok 1998). Desideratum 2 can be satisfied to reasonable accuracy if quite a large number of base-resolution pixels is used (which, however, precludes the efficient construction of simple wavelet transforms). Conversely, a tree structure seeded with a small number of base-resolution pixels forces significant variations in both the area and shape of the pixels.
- (e) The GLESP construction (Doroshkevich et al. 2005) explicitly implements the Gauss-Legendre quadrature scheme to render high accuracy in numerical integrations with respect to latitude but allows irregular variations in the pixel area and is not hierarchical—in fact, it offers no relation between the tessellations derived at different resolutions.

4. MEETING THE REQUIREMENTS: THE HEALPIX SOLUTION

All the requirements introduced in § 3 are satisfied by the class of spherical tessellations structured as follows (Górski et al. 1999).

First, let us assume that the sphere is partitioned into a number of curvilinear quadrilaterals, which constitute the base-level tessellation. If there exists a mapping of each element of partition onto a square $[0, 1] \times [0, 1]$, then a nested $n \times n$ subdivision of the square into ever diminishing subelements is obtained trivially, and a hierarchical tree structure for the resulting database follows. For example, a 2×2 partition renders a quad-

rilateral tree, which admits an elegant binary indexation (illustrated in Fig. 1) previously employed in the construction of the QuadCube spherical pixelization.

Next, let us consider the base-level spherical tessellation. An entire class of such tessellations can be constructed as illustrated in Figure 2. These constructions are characterized by two parameters: N_{θ} —the number of base-resolution pixel layers between the north and south poles and N_{ϕ} —the multiplicity of the meridional cuts, or the number of equatorial or circumpolar base-resolution pixels. Obviously, the total number of baseresolution pixels is equal to $N_{\text{base-pix}} = N_{\theta}N_{\phi}$, and the area of each one of them is equal to $\Omega_{\text{base-pix}} = 4\pi/(N_{\theta}N_{\phi})$. One should also notice that each tessellation includes two single layers of polar cap pixels (with or without an azimuthal twist in their respective positions on the sphere for odd or even values of N_{θ} , respectively) and $(N_{\theta} - 2)$ layers of equatorial zone pixels, which form a regular rhomboidal grid in the cylindrical projection of the sphere. Since the cylindrical projection is an areapreserving mapping, this property immediately illustrates that the areas of equatorial zone pixels are all equal, and to meet our requirement of a fully equal area partition of the sphere, we need to demonstrate that our constructions render identical areas for the polar pixels as well. Indeed, this allows one to formulate a constraint on the colatitude θ_* at which the lateral vertices of both polar and equatorial pixels meet:

$$2\pi(1-\cos\theta_*) = \frac{N_\phi \Omega_{\text{base-pix}}}{2}; \text{ hence } \cos\theta_* = \frac{N_\theta - 1}{N_\theta}.$$
(1)

The curvilinear quadrilateral pixels of this tessellation class retain equal areas but vary in shape, depending on their positions on the sphere. We have chosen the $N_{\theta}=3$, $N_{\phi}=4$ grid (Fig. 2, *middle row, right column*) as the definition of our digital full-sky map data standard. This choice was based on three driving requirements: that there should be no more than 4 pixels at the poles to avoid acute angles, that the elongation of equatorial pixels should be simultaneously minimized, and that the 2^n multiplicity of pixels on rings in the equatorial zone

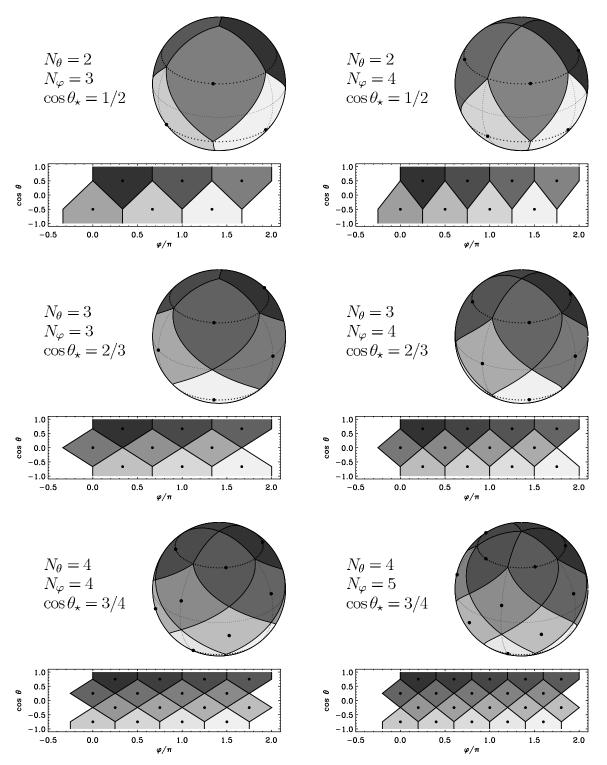


Fig. 2.—Several possible equal-area isolatitude tessellations of the sphere, which can support a hierarchical tree for the further subdivision of each large base-resolution pixel. Six variants of such a tessellation are shown here for several values of the grid parameters N_{θ} and N_{ϕ} . θ refers to the colatitude and ϕ to the longitude. For each panel two projections are shown: the top one is orthographic, the bottom one is cylindrical. θ * defines the line of constant colatitude that separates the equatorial region from the northern polar region. This is shown by the top curve in the orthographic plot; also shown is the corresponding curve separating the equatorial and southern polar regions, as well as the equator. The HEALPix implementation is described by $N_{\theta} = 3$ and $N_{\phi} = 4$.

HEALPix 763

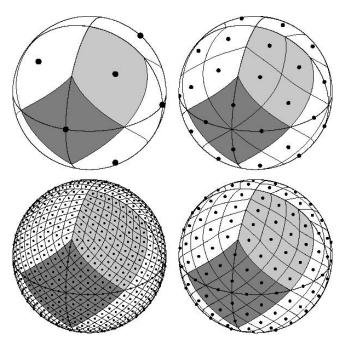


Fig. 3.—Orthographic view of the HEALPix partition of the sphere. The overplot of equator and meridians illustrates the octahedral symmetry of HEALPix. Light gray shading shows one of the 8 (4 north and 4 south) identical polar base-resolution pixels. Dark gray shading shows one of the 4 identical equatorial base-resolution pixels. Moving clockwise from the top left panel, the grid is hierarchically subdivided with the grid resolution parameter equal to $N_{\rm side}=1,\ 2,\ 4,\ 8,\$ and the corresponding total number of pixels equal to $N_{\rm pix}=12\times N_{\rm side}^2=12,\ 48,\ 192,\ 768.$ All pixel centers are located on $N_{\rm ring}=4N_{\rm side}=1$ rings of constant latitude. Within each panel the areas of all pixels are identical

should be retained for reasons related to the fast harmonic transform.

This preferred implementation, which is referred to as HEALPix, is a geometrically constructed, self-similar, refinable quadrilateral mesh on the sphere as shown in Figure 3. The base resolution comprises 12 pixels in three rings around the poles and equator. The resolution of the grid is expressed by the parameter N_{side} , which defines the number of divisions along the side of a base-resolution pixel that is needed to reach a desired high-resolution partition.¹¹ All pixel centers are placed on rings of constant latitude, and are equidistant in azimuth (on each ring). All isolatitude rings located between the upper and lower corners of the equatorial base-resolution pixels (i.e., $-\frac{2}{3}$ < $\cos \theta < \frac{2}{3}$), or in the equatorial zone, are divided into the same number of pixels: $N_{eq} = 4N_{side}$. The remaining rings are located within the polar cap regions ($|\cos \theta| > \frac{2}{3}$) and contain a varying number of pixels, increasing from ring to ring, with increasing distance from the poles, by one pixel within each quadrant. A HEALPix map has $N_{\rm pix}=12N_{\rm side}^2$ pixels of the same area $\Omega_{\rm pix}=\pi/(3N_{\rm side}^2)$.

4.1. Pixel Positions

For a resolution parameter N_{side} , the pixels are laid out on $4N_{\text{side}} - 1$ isolatitude rings.

The locations of pixel centers on the sphere are defined by $(z \equiv \cos \theta, \ \phi)$, where $\theta \in [0, \ \pi]$ is the colatitude in radians measured from the north pole and $\phi \in [0, \ 2\pi]$ is the longitude in

radians measured eastward. Pixel centers on the northern hemisphere are given by the following equations:

North polar cap.— For $p_h = (p+1)/2$, the ring index $1 \le i < N_{\text{side}}$, and the pixel-in-ring index $1 \le j \le 4i$, where

$$i = I\left(\sqrt{p_h - \sqrt{I(p_h)}}\right) + 1,\tag{2}$$

$$j = p + 1 - 2i(i - 1), \tag{3}$$

$$z = 1 - \frac{i^2}{3N_{\text{side}}^2},\tag{4}$$

$$\phi = \frac{\pi}{2i} \left(j - \frac{s}{2} \right), \quad \text{and} \quad s = 1. \tag{5}$$

North equatorial belt.—For $p'=p-2N_{\text{side}}(N_{\text{side}}-1)$, $N_{\text{side}} \leq i \leq 2N_{\text{side}}$, and $1 \leq j \leq 4N_{\text{side}}$, where

$$i = I(p'/4 N_{\text{side}}) + N_{\text{side}}, \tag{6}$$

$$j = (p' \operatorname{mod} 4 N_{\operatorname{side}}) + 1, \tag{7}$$

$$z = \frac{4}{3} - \frac{2i}{3N_{\text{side}}},\tag{8}$$

$$\phi = \frac{\pi}{2 N_{\text{side}}} \left(j - \frac{s}{2} \right), \text{ and } s = (i - N_{\text{side}} + 1) \mod 2,$$
 (9)

where the auxiliary index s describes the phase shifts along the rings and I(x) is the largest integer number smaller than x.

Pixel center positions in the southern hemisphere are obtained by the mirror symmetry of the grid with respect to the equator (z=0). One can check that the discretized area element $|\Delta z \Delta \phi| = \Omega_{\rm pix}$ is a constant by defining Δz and $\Delta \phi$ as the variation of z and ϕ when i and j, respectively, are increased by unity.

4.2. Pixel Indexing

Specific geometrical properties allow HEALPix to support two different numbering schemes for the pixels, as illustrated in Figure 4.

First, in the ring scheme, one can simply count the pixels moving down from the north to the south pole along each isolatitude ring. It is in this scheme that Fourier transforms with spherical harmonics are easy to implement. Second, one can replicate the tree structure of pixel numbering used, e.g., with the QuadCube. This can easily be implemented since, because of the simple description of pixel boundaries, the analytical mapping of the HEALPix base-resolution elements (curvilinear quadrilaterals) into a $[0,\ 1]\times[0,\ 1]$ square exists. This tree structure, aka nested scheme, allows one to implement efficiently all applications involving nearest-neighbor searches (Wandelt et al. 1998), and also allows for an immediate construction of the fast Haar wavelet transform on HEALPix.

The base-resolution pixel index number f runs in $\{0, N_{\theta}N_{\phi} - 1\} = \{0, 11\}$. Introducing the row index

$$f_{\text{row}} = I(f/N_{\phi}),\tag{10}$$

we define two functions that index the location of the southernmost corner (or vertex) of each base-resolution pixel on the sphere in latitude and longitude, respectively:

$$F_1(f) = f_{\text{row}} + 2,$$
 (11)

$$F_2(f) = 2(f \mod N_\phi) - (f_{\text{row}} \mod 2) + 1.$$
 (12)

¹¹ It should be noted that the *WMAP* team uses an alternative notation for defining various levels of resolution. Specifically, they refer to a "resolution level" defined by $N_{\rm side}=2^k$, where k can adopt the integer values $0, 1, 2, \ldots$

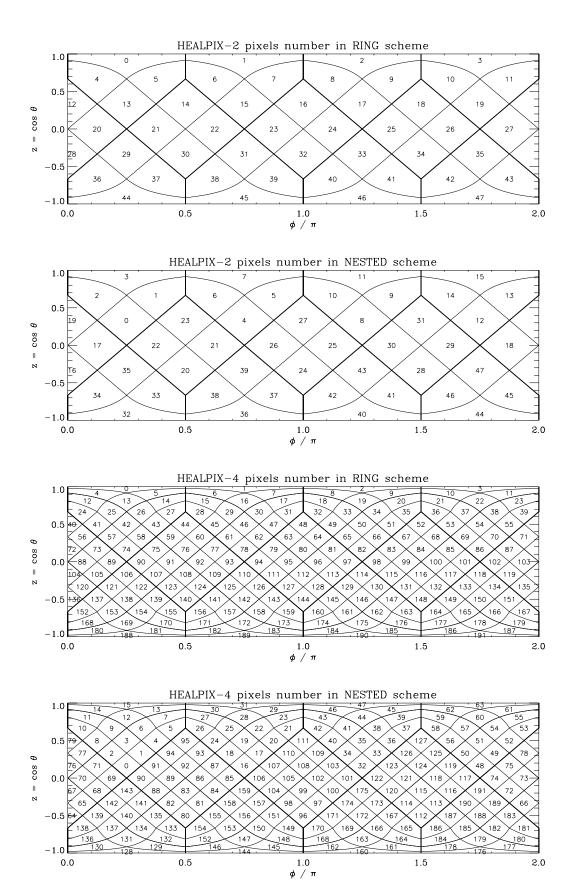


Fig. 4.—Layout of the HEALPix pixels on the sphere in a cylindrical projection and a demonstration of two possible pixel indexations—one running on isolatitude rings, the other arranged hierarchically or in a nested tree fashion. The top two panels correspond to $N_{\rm side}=2$ in first ring then nested schemes; the bottom two panels are for $N_{\rm side}=4$. θ refers to the colatitude and ϕ to the longitude.

HEALPix 765

Consider the nested index $p_n \in [0, 12 N_{\text{side}}^2 - 1]$ and define $p'_n = p_n \mod N_{\text{side}}^2$, where p'_n denotes the nested pixel index within each base-resolution element. This has the binary representation $p'_n = [\dots b_2 b_1 b_0]_2$ (and $b_i = 0$ or 1 and has weight 2^i).

Given a grid resolution parameter $N_{\rm side}$, the location of a pixel center on each base-resolution pixel is represented by the two indices x and y, both of which are members of $\{0, N_{\rm side} - 1\}$. They both have their origin in the southernmost corner of each base-resolution pixel, with the x-index running along the northeast direction, while the y-index runs along the northwest direction. The binary representation of p'_n determines the values of x and y as the following combinations of even and odd bits, respectively:

$$x = [\overline{...b_2b_0}]_2, \tag{13}$$

$$y = [\overline{...b_3b_1}]_2. \tag{14}$$

Next we introduce the vertical and horizontal coordinates (within the base-resolution pixel),

$$v = x + y, (15)$$

$$h = x - y, (16)$$

and obtain the following relation for the ring index $i \in \{1, (N_{\theta} + 1)N_{\text{side}} - 1\}$:

$$i = F_1(f)N_{\text{side}} - v - 1,$$
 (17)

as well as the longitude index

$$j = \frac{F_2(f)N_{\text{side}} + h + s}{2},$$
 (18)

which can be translated into (z, ϕ) coordinates using equations (4), (5), (8), and (9).

4.3. Pixel Boundaries

The pixel boundaries are nongeodesic and take a very simple form: $\cos\theta=a+b\phi$ in the equatorial zone and $\cos\theta=a+b/\phi^2$ in the polar caps. This allows one to explicitly check by simple analytical integration the exact equality of area among pixels and to compute efficiently more complex objects, e.g., the Fourier transforms of individual pixels.

Since the pixel center location is parameterized by the integer value of j, setting $j = k + \frac{1}{2}$ or $j = k + \frac{1}{2} + i$ (with k a positive integer in eq. [5]) and substituting into equation (4) will give for the pixel boundaries of the north polar cap $(z > \frac{2}{3})$

$$z = 1 - \frac{k^2}{3N_{\text{side}}^2} \left(\frac{\pi}{2\phi_t}\right)^2,\tag{19}$$

$$z = 1 - \frac{k^2}{3N_{\text{side}}^2} \left(\frac{\pi}{2\phi_t - \pi}\right)^2, \tag{20}$$

where $\phi_t = \phi \mod \frac{\pi}{2}$. The base pixels have boundaries defined as

$$z > \frac{2}{3}, \quad \phi = k' \frac{\pi}{2},$$
 (21)

with k' = 0, 1, 2, 3.

 $TABLE \ 1$ Summary of Number of Pixels and Pixel Sizes Accessible to HEALPix

k	$N_{\text{side}} = 2^k$	$N_{\rm pix} = 12N_{\rm side}^2$	$ heta_{ m pix} = \Omega_{ m pix}^{1/2}$
0	1	12	58°.6
1	2	48	29°.3
2	4	192	14°.7
3	8	768	7.33
4	16	3072	3°.66
5	32	12,288	1°83
6	64	49,152	55.'0
7	128	196,608	27.5
8	256	786,432	13.'7
9	512	3,145,728	6.'87
10	1024	12,582,912	3.44
11	2048	50,331,648	1.'72
12	4096	201,326,592	51."5
13	8192	805,306,368	258
14	2^{14}	3.22×10^{9}	12."9
15	2^{15}	1.29×10^{10}	6."44
16	2^{16}	5.15×10^{10}	3″22
17	2^{17}	2.06×10^{11}	1."61
:	:	:	:
29	2^{29}	3.46×10^{18}	$3\rlap.{''}93 \times 10^{-4}$

Notes.—Currently, the use of 32 bit signed integers for the pixel indexing restricts the resolution accessible to $N_{\rm side} \leq 8192$. The use of 64 bit signed integers will allow a value $N_{\rm side} = 2^{29}$ to be achieved. Note that k corresponds to the "resolution parameter" referred to by the WMAP team.

Similarly, for $\frac{2}{3} \le z \le -\frac{2}{3}$ the pixel boundaries can be found by setting $j = k + s/2 \pm (i - N_{\text{side}})/2$ in equation (9) and substituting into equation (8):

$$z = \frac{2}{3} - \frac{4k}{3N_{\text{side}}} \pm \frac{8\phi}{3\pi}.$$
 (22)

Using these pixel boundaries, one can easily check by integration that each individual pixel has the same surface area Ω_{pix} (see the Appendix).

Table 1 summarizes the number of pixels and resolutions available in HEALPix. Since all pixels have the same surface area but slightly different shape, the angular resolution is defined as

$$\theta_{\rm pix} \equiv \sqrt{\Omega_{\rm pix}}$$

$$= \sqrt{\frac{3}{\pi}} \frac{3600'}{N_{\rm cide}}.$$
(23)

4.4. The HEALPix Spherical Projection

In §§ 4.1 and 4.3 we present the algebraic description of pixel center locations and pixel boundaries of the HEALPix tessellation. In this subsection we recast these equations in the form that explicitly demonstrates the global spherical projection nature of HEALPix, i.e., the global HEALPix mapping from the surface of the sphere to the flat domain, and its inverse transformation. The existence of this transformation is essential for cartographic applications and helpful for data storage, e.g., using FITS (Calabretta & Greisen 2002). Indeed, Calabretta (2004) provides complete details on how to represent celestial coordinates for data in the HEALPix projection and subsequently store the data as a FITS image.

Consider coordinates (x_s, y_s) , where x_s runs along the longitude and y_s runs along the latitude, which are related in the

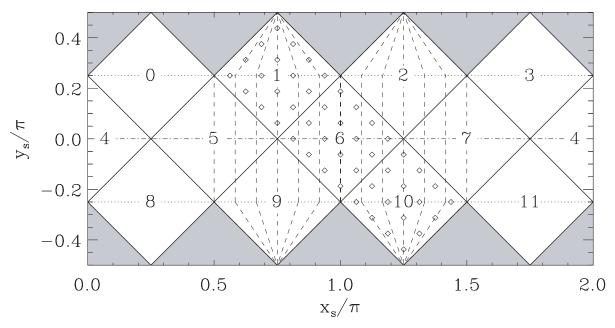


Fig. 5.—Spherical HEALPix projection onto the plane. Base-resolution HEALPix pixels (indexed here from 0 to 11, as in the HEALPix software) project into 12 identical square pixels in the plane. Hierarchical subdivision of the HEALPix grid generates identical square pixel images over the entire planar image of the HEALPix tessellation ($N_{\text{side}} = 4$ pixel centers are shown within the base pixels 1, 6, and 10). Constant-latitude lines map into horizontal lines on the plane (*dashed lines*), and the HEALPix spherical projection mapping of meridians on one hemisphere is shown by the dashed lines.

following way to the previously used HEALPix ring number $i \in \{1, 4N_{\text{side}} - 1\}$ and pixel index j (on ring i):

$$x_s = \left(j - \frac{s}{2}\right) \frac{\pi}{2N_{\text{side}}},\tag{24}$$

$$y_s = (2N_{\text{side}} - i)\frac{\pi}{4N_{\text{side}}}.$$
 (25)

Upon substituting these formulae into equations (8) and (9) (equatorial zone) and equations (19) and (20) (polar cap), we recast the pixel center positions as follows: in the HEALPix equatorial zone ($|z| < \frac{2}{3}$)

$$x_s = \phi, \tag{26}$$

$$y_s = \frac{3\pi}{8}z,\tag{27}$$

and in the HEALPix polar caps $(|z| > \frac{2}{3})$

$$x_s = \phi - (|\sigma(z)| - 1)\left(\phi_t - \frac{\pi}{4}\right),\tag{28}$$

$$y_s = \frac{\pi}{4} \sigma(z), \tag{29}$$

where $z = \cos \theta$, $\phi_t = (\phi \mod \frac{\pi}{2})$, $\sigma(z) = 2 - [3(1-z)]^{1/2}$ for z > 0, and $\sigma(-z) = -\sigma(z)$.

The same algebra renders the following representation of pixel boundaries:

$$y_s = \pm x_s + \frac{\pi}{4} - \frac{k\pi}{2N_{\text{side}}},$$
 (30)

which is now valid in both the polar and equatorial zones of HEALPix, unlike the previous description in § 4.3.

The inverse mapping from the (x_s, y_s) plane to the sphere (θ, ϕ) is given by the following equations: in the HEALPix Equatorial zone $(|y_s| < \pi/4)$

$$\phi = x_{\rm s},\tag{31}$$

$$\cos \theta = \frac{8}{3\pi} y_s,\tag{32}$$

and in the HEALPix polar caps ($|y_s| > \pi/4$)

$$\phi = x_s - \frac{|y_s| - \pi/4}{|y_s| - \pi/2} \left(x_t - \frac{\pi}{4} \right), \tag{33}$$

$$\cos \theta = \left[1 - \frac{1}{3} \left(2 - \frac{4|y_s|}{\pi}\right)^2\right] \frac{y_s}{|y_s|},\tag{34}$$

where $x_t = (x_s \mod \frac{\pi}{2})$. In the polar cap regions x_t has to satisfy the condition

$$\left| x_t - \frac{\pi}{4} \right| < \frac{\pi}{2} - |y_s|, \text{ for } \frac{\pi}{4} < |y_s| < \frac{\pi}{2}.$$
 (35)

Equations (26)–(29) and (31)–(35) provide a complete description of the HEALPix spherical projection (from the sphere to the plane, and the inverse).

The HEALPix projection of the sphere to the plane is illustrated in Figure 5. This reveals the nature of the HEALPix grid as a union of 12 identical diamond-shaped base-resolution pixels, and illustrates well the pixel area equality (all pixels of varying shape on the sphere get projected into identical square pixels on the plane). Calabretta (2004) has demonstrated that such a projection is one of a broader class, and derived the corresponding projection equations. In addition, he suggests making an extension to this representation of the HEALPix projection on a two-dimensional grid that may improve the storage of HEALPix data as FITS images.

5. SPHERICAL HARMONIC TRANSFORMS

The requirement of an isolatitude distribution for all pixel centers was built into HEALPix in order for the grid to support fast discrete spherical harmonic transforms. The reason for the fast computational time of the harmonic transform (scaling as $\sim N_{\rm nix}^{3/2}$) is entirely geometrical—the associated Legendre function components of spherical harmonics, which can only be generated via slow recursions, have to be evaluated only once for each pixel ring. For other grids that are not constrained to be isolatitude, extra computing time is wasted on the nonoptimal generation of the associated Legendre functions, which typically results in a computational performance of order $\sim N_{\rm pix}^2$. This geometrical aspect of the discrete spherical transform computation is illustrated in Figure 6, which compares HEALPix with other tessellations including the QuadCube, icosahedral tessellation of the sphere, and the ECP or "geographic grid." This plot makes it visually clear why the isolatitude ECP and HEALPix point sets support faster computation of spherical harmonic transforms than the QuadCube, the icosahedral grid, and any nonisolatitude construction.

Figure 7 demonstrates the fundamental difference between computing speeds, which can be achieved on isolatitude and nonisolatitude point sets. In order to be able to perform the necessary computational work in support of multimillion pixel spherical data sets one has to resort to isolatitude structures of point sets/sky maps, e.g., HEALPix. Moreover, future requirements are already fairly clear—the measurement of the CMB polarization will require huge multielement arrays of detectors, and will produce data sets characterized by a great multiplicity (of the order of a few thousand) of sky maps. Since there are no

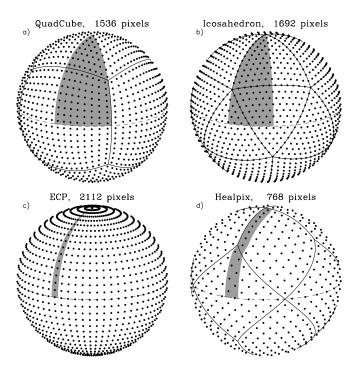


Fig. 6.—Comparison of HEALPix with other tessellations, including the QuadCube, icosahedral tessellation of the sphere, and ECP or "geographic grid." The shaded areas illustrate the subsets of all pixels on the sky for which the associated Legendre functions have to be computed in order to perform the spherical harmonic transforms. This plot demonstrates why the isolatitude ECP and HEALPix point sets support faster computation of spherical harmonic transforms than the QuadCube, the icosahedral grid, and any nonisolatitude construction.

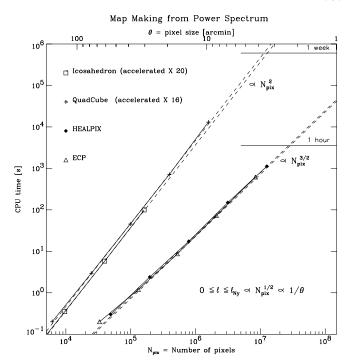


Fig. 7.—Illustration of the fundamental difference between the computing speeds that can be achieved on isolatitude and nonisolatitude point sets. In order to be able to perform the necessary computational work in support of multimillion pixel spherical data sets one has to resort to isolatitude structures of point sets/sky maps, e.g., HEALPix. Moreover, future requirements are already fairly well established—measurements of the CMB polarization will require huge multielement arrays of detectors and will produce data sets characterized by a great multiplicity (~1000) of sky maps. Since there are no computationally faster methods than those already employed in HEALPix and global synthesis/analysis of a multimillion pixel map consumes about 10³ s of CPU time on a standard serial machine, the necessary speed-up will need to be achieved via optimized parallelization of the software.

computationally faster methods than those already employed in HEALPix, and global synthesis/analysis of a multimillion pixel map consumes about 10^3 s of CPU time for a standard serial machine, the necessary speed-up will have to be achieved via optimized parallelization of the required computations.

A detailed description of the implementation and performance of spherical harmonic transforms in the HEALPix software package will be given in a separate publication.

6. SUMMARY

The Hierarchical Equal Area isoLatitude Pixelization, HEALPix, is a methodology for the discretization and fast numerical analysis and synthesis of functions or data distributed on the sphere. HEALPix is an intermediate data-structural, algorithmic, and functional layer between astronomical data, and the domain of application of a variety of science extraction tools. HEALPix as a sky map format and associated set of analysis and visualization tools has already been extensively adopted as an interface between Information Technology and Space (and suborbital) Science. This is manifested by applications of HEALPix in the following projects: CMB experiments such as the balloon-borne BOOMERANG (de Bernardis et al. 2000; Ruhl et al. 2003) and Archeops (Benoit et al. 2003a, 2003b), the satellite mission WMAP (Bennett et al. 2003), the forthcoming satellite mission *Planck*, the Sloan Digital Sky Survey, and others. Figure 8 demonstrates a selection of illustrations of some of the better known applications of HEALPix to date.

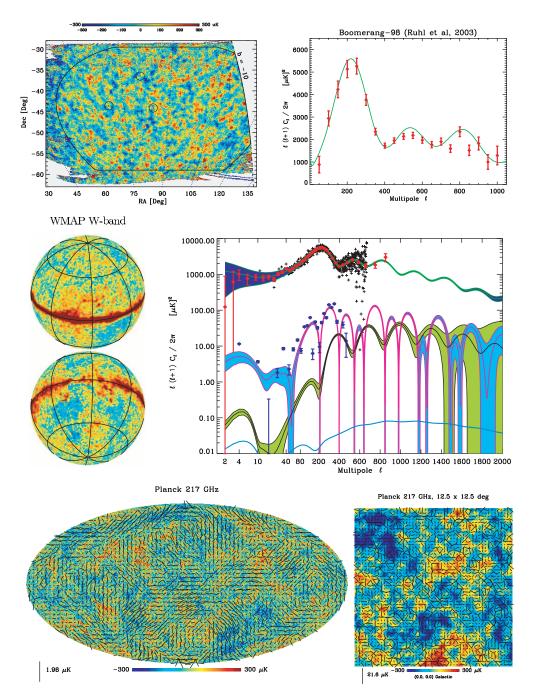


Fig. 8.—Applications of HEALPix. We show examples of both CMB data sets (existing or forthcoming) constructed using the HEALPix data structure and examples of science results derived from those data sets, specifically, power spectra of the temperature and polarization anisotropy. Sky map visualizations were generated using HEALPix tools. Top left panel: Partial sky map of the CMB sky at 150 GHz produced by the 1998 flight of the BOOMERANG experiment. Thick contour shows the area (2.9% of the sky) used for estimation of the power spectrum of the CMB anisotropy. Top right panel: Angular power spectrum of the CMB anisotropy. The red points are derived from the BOOMERANG 1998 data (within the contour on the map in the left panel) using a MASTER-like approach. The green curve shows the best fit to the first-year WMAP results. Middle left panel: Orthographic projection of the WMAP one-year W-band full-sky map smoothed with a 1° FWHM Gaussian filter. The temperature scale is the same as in the other sky map plots. Middle right panel: Comparison of current (observational) and future (predicted) determinations of the CMB power spectrum. The WMAP one-year determination of the temperature power spectrum is shown for individual multipoles with black plus signs and after binning in l as red diamonds. The signal-to-noise ratio falls to unity by l = 666. For comparison, the top green curve shows the bestfit theoretical spectrum together with the expected unbinned Planck 217 GHz measurement error. Absolute measurement values for the WMAP temperaturepolarization (TE) correlation are shown as blue diamonds (error bars divided by 10 to improve legibility). The magenta and black curves show, respectively, the WMAP best-fit TE and EE Cs. These should be contrasted with the light blue and green shaded areas, which are the predicted measurement errors expected from the Planck 217 GHz channel for a binning range of 10 in 1. Finally, the bold light blue curve shows the tensorial contribution to the polarization signal induced by gravity waves during the inflationary epoch and the lensing of the scalar polarization by large-scale structures. These signals are currently outside the range of detectability by suborbital experiments. The development of multielement detector arrays, together with space-borne experiments such as *Planck* and the Beyond-Einstein probe, will directly measure the CMB properties in the high resolution and sensitivity regime for which future HEALPix applications will be optimized. Bottom left panel: Simulation of the Planck full-sky CMB signal at 217 GHz (5' FWHM beam; this will be one of the highest angular resolution and sensitivity Planck channels) for temperature (color-coded map) and polarization (headless vectors). For clarity the polarization field has been smoothed with a 2° FWHM spatial Gaussian filter. Bottom right panel: Gnomonic projection of a 12.5 × 12.5 cutout of the full-sky map shown in the bottom left panel. This time the polarization is left unsmoothed. It is on these small (arcminute) angular scales, corresponding to the high-I regime in the spectral plot shown in the middle right panel, where future CMB experiments will probe currently unknown properties of the universe.

HEALPix 769

The development, distribution, and support of HEALPix since 1997 would have been impossible without the support of a number of institutions and individuals. We are indebted to the Theoretical Astrophysics Center in Copenhagen, Igor Novikov, and Per Rex Christensen; to the European Southern Observatory, Peter Quinn, and Kevin Maguire; to MPA, Garching, and Simon D. M. White; to Caltech's Observational Cosmology Group

and Andrew Lange; to Caltech/IPAC, George Helou, and Ken Ganga; and to JPL/Caltech, Center for Long Wavelength Astrophysics, and Charles R. Lawrence. We thank Radek Stompor for essential contributions in the early implementation of HEALPix and M. Calabretta for discussions related to § 4.4. We are also grateful for all the positive feedback received from the numerous HEALPix users worldwide.

APPENDIX

SURFACE AREA OF PIXELS

In this section, we demonstrate in detail that the boundaries introduced in § 4.3 define pixels of equal area

$$\Omega_{\rm pix} \equiv 4\pi/N_{\rm pix}.\tag{A1}$$

The calculations are performed in the cylindrical coordinate system ($z = \cos \theta$, ϕ), since the mapping from the sphere to these coordinates, (θ , ϕ) \longrightarrow (z, ϕ), is area-preserving and the pixel boundaries have a simpler form in cylindrical coordinates.

For clarity, calculations are performed for the "standard" implementation of HEALPix ($N_{\theta} = 3$, $N_{\phi} = 4$, and $N_{\text{pix}} = 12N_{\text{side}}^2$), but remain valid for any variant.

A1. EQUATORIAL REGIME

In the region $-\frac{2}{3} < z < \frac{2}{3}$ the pixel boundaries are defined in cylindrical coordinates by parallel straight lines of the form (eq. [22])

$$z_p(\phi, k) = \frac{2}{3} + ka + b\phi, \tag{A2}$$

$$z_m(\phi, k') = \frac{2}{3} + k'a - b\phi,$$
 (A3)

where $a = 4/(3N_{\text{side}})$, $b = 8/(3\pi)$, and k and k' are integers. Since the boundaries of a given pixel are parameterized by (k, k+1) and (k', k'+1) for z_p and z_m , respectively, the pixels are regular diamonds where the diagonals sit along the north-south and eastwest axes, respectively (see Fig. 2). The north-south diagonal has a length $\Delta z = a$, and the east-west one has a length $\Delta \phi = a/b$. Since the area of a diamond with orthogonal diagonals is the half product of the two diagonal lengths, the area of each pixel is

$$\Delta\Omega = \frac{1}{2}\Delta z \Delta\phi = \frac{a^2}{2b} = \frac{\pi}{3N_{\text{side}}^2} = \Omega_{\text{pix}}.$$
 (A4)

A2. POLAR REGIME

For $z > \frac{2}{3}$ and $0 \le \phi \le \pi/2$, the pixel boundaries (eq. [20]) are

$$z_p(k, \phi) = 1 - \left(\frac{kc}{\phi}\right)^2,\tag{A5}$$

$$z_m(k', \phi) = 1 - \left(\frac{k'c}{\pi/2 - \phi}\right)^2,$$
 (A6)

where $c = \pi/(12)^{1/2}/N_{\text{side}}$ and k and k' are nonnegative integers. The vertices of the pixels are the intersections of z_p and z_m for an arbitrary k, k' and are located at

$$z_v(k, k') = 1 - \frac{1}{3} \left(\frac{k + k'}{N_{\text{side}}}\right)^2,$$
 (A7)

$$\phi_v(k, k') = \frac{\pi}{2} \frac{k}{k + k'}.$$
 (A8)

As illustrated in Figure 9, each pixel, delineated by z_p and z_m with parameters k, k+1, and k', k'+1, respectively, is a curvilinear quadrilateral for which the diagonal lies along the east-west axis. This can be divided into four sectors whose area can be readily computed. The northwest sector, defined as

$$\phi < \phi_v(k, k'), \quad z > z_v(k, k'+1), \quad z < z_p(\phi, k),$$
 (A9)

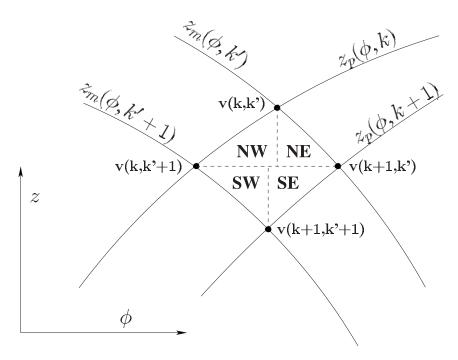


Fig. 9.—Calculation of the pixel area in the northern polar regime. z_p and z_m are the lines defining the pixel boundaries (eqs. [20] and [A6]) parameterized by the integer indices k and k', respectively, and the black circles are the pixel vertices, the location of which depend on k, k' (and N_{side}). The pixel is divided into four sectors whose area can be readily computed (see the Appendix) to prove that all pixels have the same surface area Ω_{pix} .

has a surface area of

$$\Delta\Omega_{\text{NW}} = \int_{\phi_v(k,k'+1)}^{\phi_v(k,k')} d\phi \left[z_p(\phi, k) - z_v(k, k'+1) \right] = \frac{\pi}{6N_{\text{side}}^2} \frac{k}{k+k'}. \tag{A10}$$

The northeast sector is defined as

$$\phi > \phi_v(k, k'), \quad z > z_v(k+1, k'), \quad z < z_m(\phi, k'),$$
 (A11)

which is related to the northwest sector definition (eq. [A9]) by the mapping $(k, k', \phi) \longrightarrow (k', k, \pi/2 - \phi)$. Therefore, its surface area is

$$\Delta\Omega_{\rm NE} = \frac{\pi}{6N_{\rm side}^2} \frac{k'}{k + k'}.\tag{A12}$$

The southeast sector, defined as

$$\phi > \phi_v(k+1, k'+1), \quad z < z_v(k+1, k'), \quad z > z_p(\phi, k+1),$$
(A13)

has a surface area of

$$\Delta\Omega_{\rm SE} = \int_{\phi_v(k+1,k'+1)}^{\phi_v(k+1,k')} d\phi \left[z_v(k+1,k') - z_p(\phi,k+1) \right] = \frac{\pi}{6N_{\rm side}^2} \frac{k+1}{k+k'+2}. \tag{A14}$$

The southwest sector is defined as

$$\phi < \phi_v(k+1, k'+1), \quad z < z_v(k, k'+1), \quad z > z_m(\phi, k'+1),$$
 (A15)

which is related to the southeast sector definition (eq. [A13]) by the mapping $(k, k', \phi) \longrightarrow (k', k, \pi/2 - \phi)$. Therefore,

$$\Delta\Omega_{\rm SW} = \frac{\pi}{6N_{\rm side}^2} \frac{k'+1}{k+k'+2}.\tag{A16}$$

We then find

$$\Delta\Omega = \Delta\Omega_{\rm NW} + \Delta\Omega_{\rm NE} + \Delta\Omega_{\rm SE} + \Delta\Omega_{\rm SW} = \frac{\pi}{3N_{\rm side}^2} = \Omega_{\rm pix}. \tag{A17}$$

It is worth noting that the north and south parts of the pixels have the same area

$$\Delta\Omega_{\rm NW} + \Delta\Omega_{\rm NE} = \Delta\Omega_{\rm SE} + \Delta\Omega_{\rm SW} = \frac{\Omega_{\rm pix}}{2}. \tag{A18}$$

Results obtained above for the quadrant $0 \le \phi < \pi/2$ in the north polar region can be extended to the other quadrants by translation in ϕ and to the south polar region by the mapping $z \longrightarrow -z$.

A3. TRANSITION REGIME

Pixels located at $z = \frac{2}{3}$ (resp. $z = -\frac{2}{3}$) have their north (south) boundaries defined in the polar regime (eq. [A6]) while their south (resp. north) boundaries follow the Equatorial regime (eq. [A3]). Noting that in both regimes the north and south part of the pixels have the same area of $\Omega_{pix}/2$, it is straightforward to show that all the transition pixels have the same area Ω_{pix} .

REFERENCES

Baumgardner, J. R., & Frederickson, P. O. 1985, SIAM J. Numer. Anal., 22, 1107 Bennett, C. L., et al. 2003, ApJS, 148, 1 Benoit, A., et al. 2003a, A&A, 399, L19 2003b, A&A, 399, L25 Calabretta, M. R. 2004, preprint (astro-ph/0412607) Calabretta, M. R., & Greisen, E. W. 2002, A&A, 395, 1077 Crittenden, R., & Turok, N. G. 1998, preprint (astro-ph/9806374) de Bernardis, P., et al. 2000, Nature, 404, 955 Doroshkevich, A. G., Naselsky, P. D., Verkhodanov, O. V., Novikov, D. I., Turchaninov, V. I., Novikov, I. D., Christensen, P. R., & Chiang, L.-Y. 2005, preprint (astro-ph/0501494)

Driscoll, J. R., & Healy, D. 1994, Adv. Appl. Math., 15, 202 Górski, K. M., Hivon, E., & Wandelt, B. D. 1999, in Proc. MPA/ESO Conf., Evolution of Large-Scale Structure: From Recombination to Garching, eds. A. J. Banday, R. K. Sheth, & L. Da Costa (Ipskamp: Printpartners), 37

Muciaccia, P. F, Natoli, P., & Vittorio, N. 1998, ApJ, 488, L63 Ruhl, J. E., et al. 2003, ApJ, 599, 786 Saff, E. B., & Kuijlaars, A. B. J. 1997, Math. Intel., 19, 5

Szalay, A. S., & Brunner R. J. 1999, Future Generation Comput. Syst., 16, 63 Tegmark, M. 1996, ApJ, 470, L81

Wandelt, B. D., Hivon, E., & Górski, K. M. 1998, in Fundamental Parameters in Cosmology, Proc. 33rd Rencontres de Moriond, ed. J. Tran Thanh Van (Gif-sur-Yvette: Editions Frontières), 43

White, R. A., & Stemwedel, S. W. 1992, in ASP Conf. Ser. 25, Astronomical Data Analysis, Software, and Systems I, ed. D. M. Worrall, C. Biemesderfer, & J. Barnes (San Francisco: ASP), 379